



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Yokoi, Haruhiko  
Shiotsu, Yukimasa  
Konishi, Noboru
- (ii) TITLE OF INVENTION: NOVEL POLYPEPTIDES
- (iii) NUMBER OF SEQUENCES: 46
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: NIXON & VANDERHYE P.C.
  - (B) STREET: 1100 North Glebe Rd. 8th floor
  - (C) CITY: Arlington
  - (D) STATE: VA
  - (E) COUNTRY: USA
  - (F) ZIP: 22201-4741
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 09/680,514
  - (B) FILING DATE: 6-OCT-2000
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/765,337
  - (B) FILING DATE: 23-DEC-1996
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: PCT/JP96/01157
  - (B) FILING DATE: 26-APR-1996
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: JP P.HEI.7-102625
  - (B) FILING DATE: 26-APR-1995
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Crawford, Arthur R.
  - (B) REGISTRATION NUMBER: 25,327
  - (C) REFERENCE/DOCKET NUMBER: 249-89
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 703-816-4000
  - (B) TELEFAX: 703-816-4100

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 328 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val	Leu	Ser	Lys	Leu	Leu	1	5	10	15
Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser	Gln	Cys	Pro	Glu	Val	20	25	30	
His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala	Val	Asp	Phe	Ser	Leu	35	40	45	
Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys	Ala	Gln	Asp	Ile	Leu	50	55	60	
Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met	Ala	Ala	Arg	Gly	Gln	65	70	75	80
Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly	Gln	Leu	Ser	Gly	Gln	85	90	95	
Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu	Leu	Gly	Thr	Gln	Leu	100	105	110	
Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp	Pro	Asn	Ala	Ile	Phe	115	120	125	
Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val	Arg	Phe	Leu	Met	Leu	130	135	140	
Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Arg	Ala	Pro	Thr	Tyr	Arg	Ala	145	150	155	160
Ser	Ser	Leu	Pro	Gln	Ser	Phe	Leu	Leu	Lys	Ser	Leu	Glu	Gln	Val	Arg	165	170	175	
Lys	Ile	Gln	Gly	Asp	Gly	Ala	Ala	Leu	Gln	Glu	Lys	Leu	Cys	Ala	Thr	180	185	190	
Tyr	Lys	Leu	Cys	His	Pro	Glu	Glu	Leu	Val	Leu	Leu	Gly	His	Ser	Leu	195	200	205	
Gly	Ile	Pro	Trp	Ala	Pro	Leu	Ser	Ser	Cys	Pro	Ser	Gln	Ala	Leu	Gln	210	215	220	
Leu	Ala	Gly	Cys	Leu	Ser	Gln	Leu	His	Ser	Gly	Leu	Phe	Leu	Tyr	Gln	225	230	235	240

Gly	Leu	Leu	Gln	Ala	Leu	Glu	Gly	Ile	Ser	Pro	Glu	Leu	Gly	Pro	Thr	245	250	255
Leu	Asp	Thr	Leu	Gln	Leu	Asp	Val	Ala	Asp	Phe	Ala	Thr	Thr	Ile	Trp	260	265	270
Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala	Pro	Ala	Leu	Gln	Pro	Thr	Gln	275	280	285
Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala	Phe	Gln	Arg	Arg	Ala	Gly	Gly	290	295	300
Val	Leu	Val	Ala	Ser	His	Leu	Gln	Ser	Phe	Leu	Glu	Val	Ser	Tyr	Arg	305	310	315
Val	Leu	Arg	His	Leu	Ala	Gln	Pro									325		

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 340 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val	Leu	Ser	Lys	Leu	Leu	1	5	10	15
Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser	Gln	Cys	Pro	Glu	Val	20	25	30	
His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala	Val	Asp	Phe	Ser	Leu	35	40	45	
Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys	Ala	Gln	Asp	Ile	Leu	50	55	60	
Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met	Ala	Ala	Arg	Gly	Gln	65	70	75	80
Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly	Gln	Leu	Ser	Gly	Gln	85	90	95	
Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu	Leu	Gly	Thr	Gln	Leu	100	105	110	
Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp	Pro	Asn	Ala	Ile	Phe	115	120	125	

Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val	Arg	Phe	Leu	Met	Leu
130							135					140			
Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Gly	Gly	Gly	Ser	Gly	Gly	Gly
145					150					155					160
Ser	Gly	Gly	Gly	Ser	Arg	Ala	Pro	Thr	Tyr	Arg	Ala	Ser	Ser	Leu	Pro
				165					170					175	
Gln	Ser	Phe	Leu	Leu	Lys	Ser	Leu	Glu	Gln	Val	Arg	Lys	Ile	Gln	Gly
			180					185					190		
Asp	Gly	Ala	Ala	Leu	Gln	Glu	Lys	Leu	Cys	Ala	Thr	Tyr	Lys	Leu	Cys
		195					200					205			
His	Pro	Glu	Glu	Leu	Val	Leu	Leu	Gly	His	Ser	Leu	Gly	Ile	Pro	Trp
	210					215					220				
Ala	Pro	Leu	Ser	Ser	Cys	Pro	Ser	Gln	Ala	Leu	Gln	Leu	Ala	Gly	Cys
225					230					235					240
Leu	Ser	Gln	Leu	His	Ser	Gly	Leu	Phe	Leu	Tyr	Gln	Gly	Leu	Leu	Gln
				245					250					255	
Ala	Leu	Glu	Gly	Ile	Ser	Pro	Glu	Leu	Gly	Pro	Thr	Leu	Asp	Thr	Leu
			260					265					270		
Gln	Leu	Asp	Val	Ala	Asp	Phe	Ala	Thr	Thr	Ile	Trp	Gln	Gln	Met	Glu
		275					280					285			
Glu	Leu	Gly	Met	Ala	Pro	Ala	Leu	Gln	Pro	Thr	Gln	Gly	Ala	Met	Pro
	290					295					300				
Ala	Phe	Ala	Ser	Ala	Phe	Gln	Arg	Arg	Ala	Gly	Gly	Val	Leu	Val	Ala
305					310					315					320
Ser	His	Leu	Gln	Ser	Phe	Leu	Glu	Val	Ser	Tyr	Arg	Val	Leu	Arg	His
				325					330					335	
Leu	Ala	Gln	Pro												
			340												

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val	Leu	Ser	Lys	Leu	Leu	1	5	10	15
Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser	Gln	Cys	Pro	Glu	Val	20	25	30	
His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala	Val	Asp	Phe	Ser	Leu	35	40	45	
Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys	Ala	Gln	Asp	Ile	Leu	50	55	60	
Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met	Ala	Ala	Arg	Gly	Gln	65	70	75	80
Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly	Gln	Leu	Ser	Gly	Gln	85	90	95	
Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu	Leu	Gly	Thr	Gln	Leu	100	105	110	
Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp	Pro	Asn	Ala	Ile	Phe	115	120	125	
Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val	Arg	Phe	Leu	Met	Leu	130	135	140	
Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Ser	Gly	Gly	Gly	Ser	Gly	Gly	145	150	155	160
Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Arg	Ala	Pro	Thr	Tyr	Arg	Ala	165	170	175	
Ser	Ser	Leu	Pro	Gln	Ser	Phe	Leu	Leu	Lys	Ser	Leu	Glu	Gln	Val	Arg	180	185	190	
Lys	Ile	Gln	Gly	Asp	Gly	Ala	Ala	Leu	Gln	Glu	Lys	Leu	Cys	Ala	Thr	195	200	205	
Tyr	Lys	Leu	Cys	His	Pro	Glu	Glu	Leu	Val	Leu	Leu	Gly	His	Ser	Leu	210	215	220	
Gly	Ile	Pro	Trp	Ala	Pro	Leu	Ser	Ser	Cys	Pro	Ser	Gln	Ala	Leu	Gln	225	230	235	240
Leu	Ala	Gly	Cys	Leu	Ser	Gln	Leu	His	Ser	Gly	Leu	Phe	Leu	Tyr	Gln	245	250	255	
Gly	Leu	Leu	Gln	Ala	Leu	Glu	Gly	Ile	Ser	Pro	Glu	Leu	Gly	Pro	Thr	260	265	270	
Leu	Asp	Thr	Leu	Gln	Leu	Asp	Val	Ala	Asp	Phe	Ala	Thr	Thr	Ile	Trp	275	280	285	
Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala	Pro	Ala	Leu	Gln	Pro	Thr	Gln	290	295	300	

Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln Arg Arg Ala Gly Gly  
 305 310 315 320

Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr Arg  
 325 330 335

Val Leu Arg His Leu Ala Gln Pro  
 340

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1047 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 1..63

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 64..1047

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1047

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATG GAG CTG ACT GAA TTG CTC CTC GTG GTC ATG CTT CTC CTA ACT GCA	48
Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala	
-21 -20 -15 -10	
AGG CTA ACG CTG TCC AGC CCG GCT CCT CCT GCT TGT GAC CTC CGA GTC	96
Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val	
-5 1 5 10	
CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT CAC AGC AGA CTG AGC	144
Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser	
15 20 25	
CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA CCT GTC CTG CTG CCT GCT	192
Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala	
30 35 40	
GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG	240
Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys	
45 50 55	

GCA CAG GAC ATT CTG GGA GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG	288
Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met	
60 65 70 75	
GCA GCA CGG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG	336
Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly	
80 85 90	
CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC	384
Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu	
95 100 105	
CTT GGA ACC CAG CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT	432
Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp	
110 115 120	
CCC AAT GCC ATC TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG	480
Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val	
125 130 135	
CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTA CGG CGG GCG	528
Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala	
140 145 150 155	
CCA ACA TAT CGC GCC TCG AGT CTA CCA CAG AGC TTC CTT TTA AAA AGC	576
Pro Thr Tyr Arg Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser	
160 165 170	
TTA GAG CAA GTG AGG AAG ATC CAG GGC GAT GGC GCA GCG CTC CAG GAG	624
Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu	
175 180 185	
AAG CTG TGT GCC ACC TAC AAG CTG TGC CAC CCC GAG GAG CTG GTG CTG	672
Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu	
190 195 200	
CTC GGA CAC TCT CTG GGC ATC CCC TGG GCT CCC CTG AGC AGC TGC CCC	720
Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro	
205 210 215	
AGC CAG GCC CTG CAG CTG GCA GGC TGC TTG AGC CAA CTC CAT AGC GGC	768
Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly	
220 225 230 235	
CTT TTC CTC TAC CAG GGG CTC CTG CAG GCC CTG GAA GGG ATC TCC CCC	816
Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro	
240 245 250	
GAG TTG GGT CCC ACC TTG GAC ACA CTG CAG CTG GAC GTC GCC GAC TTT	864
Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe	
255 260 265	
GCC ACC ACC ATC TGG CAG CAG ATG GAA GAA CTG GGA ATG GCC CCT GCC	912
Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala	
270 275 280	
CTG CAG CCC ACC CAG GGT GCC ATG CCG GCC TTC GCC TCT GCT TTC CAG	960

Leu	Gln	Pro	Thr	Gln	Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala	Phe	Gln		
285						290					295						
CGC	CGG	GCA	GGA	GGG	GTC	CTA	GTT	GCC	TCC	CAT	CTG	CAG	AGC	TTC	CTG	1008	
Arg	Arg	Ala	Gly	Gly	Val	Leu	Val	Ala	Ser	His	Leu	Gln	Ser	Phe	Leu		
300					305					310					315		
GAG	GTG	TCG	TAC	CGC	GTT	CTA	CGC	CAC	CTT	GCC	CAG	CCC				1047	
Glu	Val	Ser	Tyr	Arg	Val	Leu	Arg	His	Leu	Ala	Gln	Pro					
				320					325								

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Glu	Leu	Thr	Glu	Leu	Leu	Leu	Val	Val	Met	Leu	Leu	Leu	Thr	Ala		
-21	-20					-15				-10							
Arg	Leu	Thr	Leu	Ser	Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val		
-5					1				5					10			
Leu	Ser	Lys	Leu	Leu	Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser		
			15					20					25				
Gln	Cys	Pro	Glu	Val	His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala		
		30					35					40					
Val	Asp	Phe	Ser	Leu	Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys		
	45					50					55						
Ala	Gln	Asp	Ile	Leu	Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met		
60					65					70					75		
Ala	Ala	Arg	Gly	Gln	Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly		
				80					85					90			
Gln	Leu	Ser	Gly	Gln	Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu		
			95				100						105				
Leu	Gly	Thr	Gln	Leu	Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp		
	110						115					120					
Pro	Asn	Ala	Ile	Phe	Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val		
	125					130					135						
Arg	Phe	Leu	Met	Leu	Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Arg	Ala		
140					145					150					155		



Pro Thr Tyr Arg Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser  
 160 165 170  
 Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu  
 175 180 185  
 Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu  
 190 195 200  
 Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro  
 205 210 215  
 Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly  
 220 225 230 235  
 Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro  
 240 245 250  
 Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe  
 255 260 265  
 Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala  
 270 275 280  
 Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln  
 285 290 295  
 Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu  
 300 305 310 315  
 Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro  
 320 325

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 1..63

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 64..1083

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1083

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG GAG CTG ACT GAA TTG CTC CTC GTG GTC ATG CTT CTC CTA ACT GCA Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala -21 -20 -15 -10	48
AGG CTA ACG CTG TCC AGC CCG GCT CCT CCT GCT TGT GAC CTC CGA GTC Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val -5 1 5 10	96
CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT CAC AGC AGA CTG AGC Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser 15 20 25	144
CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA CCT GTC CTG CTG CCT GCT Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala 30 35 40	192
GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys 45 50 55	240
GCA CAG GAC ATT CTG GGA GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met 60 65 70 75	288
GCA GCA CGG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly 80 85 90	336
CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu 95 100 105	384
CTT GGA ACC CAG CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp 110 115 120	432
CCC AAT GCC ATC TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val 125 130 135	480
CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTC AGG GGT GGC Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Gly Gly 140 145 150 155	528
GGT TCT GGA GGT GGT TCC GGA GGG GGT TCT AGA GCA CCA ACA TAT CGC Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Arg Ala Pro Thr Tyr Arg 160 165 170	576
GCC TCG AGT CTA CCA CAG AGC TTC CTT TTA AAA AGC TTA GAG CAA GTG Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser Leu Glu Gln Val 175 180 185	624
AGG AAG ATC CAG GGC GAT GGC GCA GCG CTC CAG GAG AAG CTG TGT GCC	672

Arg	Lys	Ile	Gln	Gly	Asp	Gly	Ala	Ala	Leu	Gln	Glu	Lys	Leu	Cys	Ala	
		190					195					200				
ACC	TAC	AAG	CTG	TGC	CAC	CCC	GAG	GAG	CTG	GTG	CTG	CTC	GGA	CAC	TCT	720
Thr	Tyr	Lys	Leu	Cys	His	Pro	Glu	Glu	Leu	Val	Leu	Leu	Gly	His	Ser	
	205					210					215					
CTG	GGC	ATC	CCC	TGG	GCT	CCC	CTG	AGC	AGC	TGC	CCC	AGC	CAG	GCC	CTG	768
Leu	Gly	Ile	Pro	Trp	Ala	Pro	Leu	Ser	Ser	Cys	Pro	Ser	Gln	Ala	Leu	
220					225					230					235	
CAG	CTG	GCA	GGC	TGC	TTG	AGC	CAA	CTC	CAT	AGC	GGC	CTT	TTC	CTC	TAC	816
Gln	Leu	Ala	Gly	Cys	Leu	Ser	Gln	Leu	His	Ser	Gly	Leu	Phe	Leu	Tyr	
				240					245					250		
CAG	GGG	CTC	CTG	CAG	GCC	CTG	GAA	GGG	ATC	TCC	CCC	GAG	TTG	GGT	CCC	864
Gln	Gly	Leu	Leu	Gln	Ala	Leu	Glu	Gly	Ile	Ser	Pro	Glu	Leu	Gly	Pro	
			255					260					265			
ACC	TTG	GAC	ACA	CTG	CAG	CTG	GAC	GTC	GCC	GAC	TTT	GCC	ACC	ACC	ATC	912
Thr	Leu	Asp	Thr	Leu	Gln	Leu	Asp	Val	Ala	Asp	Phe	Ala	Thr	Thr	Ile	
		270					275					280				
TGG	CAG	CAG	ATG	GAA	GAA	CTG	GGA	ATG	GCC	CCT	GCC	CTG	CAG	CCC	ACC	960
Trp	Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala	Pro	Ala	Leu	Gln	Pro	Thr	
	285					290					295					
CAG	GGT	GCC	ATG	CCG	GCC	TTC	GCC	TCT	GCT	TTC	CAG	CGC	CGG	GCA	GGA	1008
Gln	Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala	Phe	Gln	Arg	Arg	Ala	Gly	
300					305					310					315	
GGG	GTC	CTA	GTT	GCC	TCC	CAT	CTG	CAG	AGC	TTC	CTG	GAG	GTG	TCG	TAC	1056
Gly	Val	Leu	Val	Ala	Ser	His	Leu	Gln	Ser	Phe	Leu	Glu	Val	Ser	Tyr	
				320					325					330		
CGC	GTT	CTA	CGC	CAC	CTT	GCC	CAG	CCC								1083
Arg	Val	Leu	Arg	His	Leu	Ala	Gln	Pro								
			335				340									

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 361 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Glu	Leu	Thr	Glu	Leu	Leu	Leu	Val	Val	Met	Leu	Leu	Leu	Thr	Ala
-21	-20						-15				-10				
Arg	Leu	Thr	Leu	Ser	Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val
-5					1				5					10	

Leu	Ser	Lys	Leu	Leu	Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser	15	20	25
Gln	Cys	Pro	Glu	Val	His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala	30	35	40
Val	Asp	Phe	Ser	Leu	Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys	45	50	55
Ala	Gln	Asp	Ile	Leu	Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met	60	65	70
Ala	Ala	Arg	Gly	Gln	Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly	80	85	90
Gln	Leu	Ser	Gly	Gln	Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu	95	100	105
Leu	Gly	Thr	Gln	Leu	Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp	110	115	120
Pro	Asn	Ala	Ile	Phe	Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val	125	130	135
Arg	Phe	Leu	Met	Leu	Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Gly	Gly	140	145	150
Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Ser	Arg	Ala	Pro	Thr	Tyr	Arg	160	165	170
Ala	Ser	Ser	Leu	Pro	Gln	Ser	Phe	Leu	Leu	Lys	Ser	Leu	Glu	Gln	Val	175	180	185
Arg	Lys	Ile	Gln	Gly	Asp	Gly	Ala	Ala	Leu	Gln	Glu	Lys	Leu	Cys	Ala	190	195	200
Thr	Tyr	Lys	Leu	Cys	His	Pro	Glu	Glu	Leu	Val	Leu	Leu	Gly	His	Ser	205	210	215
Leu	Gly	Ile	Pro	Trp	Ala	Pro	Leu	Ser	Ser	Cys	Pro	Ser	Gln	Ala	Leu	220	225	230
Gln	Leu	Ala	Gly	Cys	Leu	Ser	Gln	Leu	His	Ser	Gly	Leu	Phe	Leu	Tyr	240	245	250
Gln	Gly	Leu	Leu	Gln	Ala	Leu	Glu	Gly	Ile	Ser	Pro	Glu	Leu	Gly	Pro	255	260	265
Thr	Leu	Asp	Thr	Leu	Gln	Leu	Asp	Val	Ala	Asp	Phe	Ala	Thr	Thr	Ile	270	275	280
Trp	Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala	Pro	Ala	Leu	Gln	Pro	Thr	285	290	295
Gln	Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala	Phe	Gln	Arg	Arg	Ala	Gly	300	305	310
																		315

Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr  
320 325 330

Arg Val Leu Arg His Leu Ala Gln Pro  
335 340

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1095 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 64..1095

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1095

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATG GAG CTG ACT GAA TTG CTC CTC GTG GTC ATG CTT CTC CTA ACT GCA Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala -21 -20 -15 -10	48
AGG CTA ACG CTG TCC AGC CCG GCT CCT CCT GCT TGT GAC CTC CGA GTC Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val -5 1 5 10	96
CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT CAC AGC AGA CTG AGC Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser 15 20 25	144
CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA CCT GTC CTG CTG CCT GCT Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala 30 35 40	192
GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys 45 50 55	240
GCA CAG GAC ATT CTG GGA GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met 60 65 70 75	288
GCA GCA CGG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly 80 85 90	336

CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu 95 100 105	384
CTT GGA ACC CAG CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp 110 115 120	432
CCC AAT GCC ATC TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val 125 130 135	480
CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTA CGG TCC GGA Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Ser Gly 140 145 150 155	528
GGT GGC TCT GGC GGT GGT TCT GGT GGC GGC TCC GGA GGC GGT CGT GCG Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Arg Ala 160 165 170	576
CCA ACA TAT CGC GCC TCG AGT CTA CCA CAG AGC TTC CTT TTA AAA AGC Pro Thr Tyr Arg Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser 175 180 185	624
TTA GAG CAA GTG AGG AAG ATC CAG GGC GAT GGC GCA GCG CTC CAG GAG Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu 190 195 200	672
AAG CTG TGT GCC ACC TAC AAG CTG TGC CAC CCC GAG GAG CTG GTG CTG Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu 205 210 215	720
CTC GGA CAC TCT CTG GGC ATC CCC TGG GCT CCC CTG AGC AGC TGC CCC Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro 220 225 230 235	768
AGC CAG GCC CTG CAG CTG GCA GGC TGC TTG AGC CAA CTC CAT AGC GGC Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly 240 245 250	816
CTT TTC CTC TAC CAG GGG CTC CTG CAG GCC CTG GAA GGG ATC TCC CCC Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro 255 260 265	864
GAG TTG GGT CCC ACC TTG GAC ACA CTG CAG CTG GAC GTC GCC GAC TTT Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe 270 275 280	912
GCC ACC ACC ATC TGG CAG CAG ATG GAA GAA CTG GGA ATG GCC CCT GCC Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala 285 290 295	960
CTG CAG CCC ACC CAG GGT GCC ATG CCG GCC TTC GCC TCT GCT TTC CAG Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln 300 305 310 315	1008
CGC CGG GCA GGA GGG GTC CTA GTT GCC TCC CAT CTG CAG AGC TTC CTG	1056

Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu  
320 325 330

GAG GTG TCG TAC CGC GTT CTA CGC CAC CTT GCC CAG CCC  
Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro  
335 340

1095

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 365 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala  
-21 -20 -15 -10

Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val  
-5 1 5 10

Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser  
15 20 25

Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala  
30 35 40

Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys  
45 50 55

Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met  
60 65 70 75

Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly  
80 85 90

Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu  
95 100 105

Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp  
110 115 120

Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val  
125 130 135

Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Ser Gly  
140 145 150 155

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Arg Ala  
160 165 170

Pro Thr Tyr Arg Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser  
175 180 185

Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu  
 190 195 200  
 Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu  
 205 210 215  
 Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro  
 220 225 230 235  
 Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly  
 240 245 250  
 Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro  
 255 260 265  
 Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe  
 270 275 280  
 Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala  
 285 290 295  
 Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln  
 300 305 310 315  
 Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu  
 320 325 330  
 Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro  
 335 340

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 27..44

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTCTCCAAGC TTGAATTCG GCCAGAATGG AGCTGACTGA ATTG

44

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid



- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GTAGAGGTAC CGCGGCCGCT TACCCTTCCT GAGACAGATT CTGGGAG

47

- (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TGAACCTCTG GGCACCTGGCT CAGT

24

- (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCTGCCTGCT GTGGACTTTA GCTT

24

- (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGT TGG AAG CTC AGG AAG ATG GCA  
Cys Trp Lys Leu Arg Lys Met Ala  
1 5

24

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Cys Trp Lys Leu Arg Lys Met Ala  
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 1..24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCT GAT GCT TGT AGG AGG GTC CAC  
Pro Asp Ala Cys Arg Arg Val His  
1 5

24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Pro Asp Ala Cys Arg Arg Val His  
1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TCA AGA GTT CGT GTA TCC TGT TCA  
Ser Arg Val Arg Val Ser Cys Ser  
1 5

24

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ser Arg Val Arg Val Ser Cys Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GAA TGG AAC TCG TGG ACT CTT TCC  
 Glu Trp Asn Ser Trp Thr Leu Ser  
   1                  5

24

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Glu Trp Asn Ser Trp Thr Leu Ser  
   1                  5

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTAAAACGAC GGCCAGT

17

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CAGGAAACAG CTATGAC

17

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 66 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: join(1..3, 43..66)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TGC TCTAGAACCG CCTCCGGAAC CACCTCCAGA ACCGCCACC CCT GAC GCA GAG  
Cys Pro Asp Ala Glu  
1 5

54

GGT GGA CCC TCC  
Gly Gly Pro Ser

66

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Cys Pro Asp Ala Glu Gly Gly Pro Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 45 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 22..45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGTTCGGAG GCGGTTCTAG A GCA CCA ACA TAT CGC GCC TCG AGT  
Ala Pro Thr Tyr Arg Ala Ser Ser  
1 5

45

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ala Pro Thr Tyr Arg Ala Ser Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CATTCCGCGG GGTACCGCGG CCGCTCAGGG CTGGGCAAGG TGGCGTAG

48

- (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..24

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGC TGC TTG AGC CAA CTC CAT AGC  
Gly Cys Leu Ser Gln Leu His Ser  
1 5

24

- (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Gly Cys Leu Ser Gln Leu His Ser  
1 5

- (2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..24

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GAC CCA ACT CGG GGG AGA TCC CTT  
 Asp Pro Thr Arg Gly Arg Ser Leu  
   1                  5

24

- (2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Asp Pro Thr Arg Gly Arg Ser Leu  
   1                  5

- (2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 57 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (ix) FEATURE:
  - (A) NAME/KEY: mutation
  - (B) LOCATION: replace(25, "")

- (ix) FEATURE:
  - (A) NAME/KEY: mutation
  - (B) LOCATION: replace(33..34, "")

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:



## (2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly  
1                      5                      10                      15

Arg

## (2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 61 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..6

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 7..57
- (D) OTHER INFORMATION: /product= "linker peptide"

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..5
- (D) OTHER INFORMATION: /product= "SplI"

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 7..12
- (D) OTHER INFORMATION: /product= "MroI"

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature

(B) LOCATION: 43..48  
(D) OTHER INFORMATION: /product= "MroI"

(ix) FEATURE:  
  (A) NAME/KEY: misc\_feature  
  (B) LOCATION: 58..61  
  (D) OTHER INFORMATION: /product= "BbeI"

(ix) FEATURE:  
  (A) NAME/KEY: mutation  
  (B) LOCATION: replace(4..5, "")

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GTA CGG TCCGGAGGTG GCTCTGGCGG TGGTTCCTGGT GGCGGCTCCG GAGGCGGTCG	56
Val Arg	
1	
TGCGC	61

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
  (A) LENGTH: 2 amino acids  
  (B) TYPE: amino acid  
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Val Arg  
1

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:  
  (A) LENGTH: 53 base pairs  
  (B) TYPE: nucleic acid  
  (C) STRANDEDNESS: single  
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
  (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:  
  (A) NAME/KEY: misc\_feature  
  (B) LOCATION: 1..51  
  (D) OTHER INFORMATION: /product= "linker peptide"

(ix) FEATURE:  
  (A) NAME/KEY: misc\_feature  
  (B) LOCATION: 53

(D) OTHER INFORMATION: /product= "SplI"

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 10..15

(D) OTHER INFORMATION: /product= "MroI"

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 46..51

(D) OTHER INFORMATION: /product= "MroI"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

ACGACCGCCT CCGGAGCCGC CACCAGAACC ACCGCCAGAG CCACCTCCGG ACC

53

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 175 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1

(D) OTHER INFORMATION: Xaa = Met

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Xaa Thr Pro Leu Gly Pro Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu  
1 5 10 15

Lys Cys Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu  
20 25 30

Gln Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Gly Leu  
35 40 45

Val Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser  
50 55 60

Cys Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His  
65 70 75 80

Ser Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile  
85 90 95

Ser Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala  
100 105 110

Asp	Phe	Ala	Thr	Thr	Ile	Trp	Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala
		115					120					125			
Pro	Ala	Leu	Gln	Pro	Thr	Gln	Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala
		130				135					140				
Phe	Gln	Arg	Arg	Ala	Gly	Gly	Val	Leu	Val	Ala	Ser	His	Leu	Gln	Ser
145					150					155				160	
Phe	Leu	Glu	Val	Ser	Tyr	Arg	Val	Leu	Arg	His	Leu	Ala	Gln	Pro	
				165					170					175	

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 332 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ser	Pro	Ala	Pro	Pro	Ala	Cys	Asp	Leu	Arg	Val	Leu	Ser	Lys	Leu	Leu
1				5					10					15	
Arg	Asp	Ser	His	Val	Leu	His	Ser	Arg	Leu	Ser	Gln	Cys	Pro	Glu	Val
			20					25					30		
His	Pro	Leu	Pro	Thr	Pro	Val	Leu	Leu	Pro	Ala	Val	Asp	Phe	Ser	Leu
		35					40					45			
Gly	Glu	Trp	Lys	Thr	Gln	Met	Glu	Glu	Thr	Lys	Ala	Gln	Asp	Ile	Leu
	50					55					60				
Gly	Ala	Val	Thr	Leu	Leu	Leu	Glu	Gly	Val	Met	Ala	Ala	Arg	Gly	Gln
65					70					75				80	
Leu	Gly	Pro	Thr	Cys	Leu	Ser	Ser	Leu	Leu	Gly	Gln	Leu	Ser	Gly	Gln
				85					90					95	
Val	Arg	Leu	Leu	Leu	Gly	Ala	Leu	Gln	Ser	Leu	Leu	Gly	Thr	Gln	Leu
			100					105					110		
Pro	Pro	Gln	Gly	Arg	Thr	Thr	Ala	His	Lys	Asp	Pro	Asn	Ala	Ile	Phe
		115					120					125			
Leu	Ser	Phe	Gln	His	Leu	Leu	Arg	Gly	Lys	Val	Arg	Phe	Leu	Met	Leu
	130					135					140				
Val	Gly	Gly	Ser	Thr	Leu	Cys	Val	Arg	Arg	Ala	Pro	Pro	Thr	Thr	Ala

145		150		155		160									
Val	Pro	Ser	Arg	Thr	Ser	Leu	Val	Leu	Thr	Leu	Asn	Glu	Leu	Pro	Asn
				165					170					175	
Arg	Thr	Ser	Gly	Leu	Leu	Glu	Thr	Asn	Phe	Thr	Ala	Ser	Ala	Arg	Thr
			180					185					190		
Thr	Gly	Ser	Gly	Leu	Leu	Lys	Trp	Gln	Gln	Gly	Phe	Arg	Ala	Lys	Ile
		195					200					205			
Pro	Gly	Leu	Leu	Asn	Gln	Thr	Ser	Arg	Ser	Leu	Asp	Gln	Ile	Pro	Gly
	210					215					220				
Tyr	Leu	Asn	Arg	Ile	His	Glu	Leu	Leu	Asn	Gly	Thr	Arg	Gly	Leu	Phe
225					230					235					240
Pro	Gly	Pro	Ser	Arg	Arg	Thr	Leu	Gly	Ala	Pro	Asp	Ile	Ser	Ser	Gly
				245					250					255	
Thr	Ser	Asp	Thr	Gly	Ser	Leu	Pro	Pro	Asn	Leu	Gln	Pro	Gly	Tyr	Ser
			260					265					270		
Pro	Ser	Pro	Thr	His	Pro	Pro	Thr	Gly	Gln	Tyr	Thr	Leu	Phe	Pro	Leu
		275					280					285			
Pro	Pro	Thr	Leu	Pro	Thr	Pro	Val	Val	Gln	Leu	His	Pro	Leu	Leu	Pro
	290					295					300				
Asp	Pro	Ser	Ala	Pro	Thr	Pro	Thr	Pro	Thr	Ser	Pro	Leu	Leu	Asn	Thr
305					310					315					320
Ser	Tyr	Thr	His	Ser	Gln	Asn	Leu	Ser	Gln	Glu	Gly				
			325						330						

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ser	Gly	Gly	Gly	Arg
1				5

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly  
1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO:43:

- (i) \*SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Ser Gly Gly Gly  
1

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
1                      5                      10

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Gly Gly Gly Ser Gly Gly Gly Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Thr Pro Leu Gly Pro Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys  
1 5 10 15

Cys Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln  
20 25 30

Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Gly Leu Val  
35 40 45

Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys  
50 55 60

Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser  
65 70 75 80

Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser  
85 90 95

Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala Asp  
100 105 110

Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala Pro  
115 120 125

Ala Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe  
130 135 140

Gln Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe  
145 150 155 160

Leu Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro  
165 170